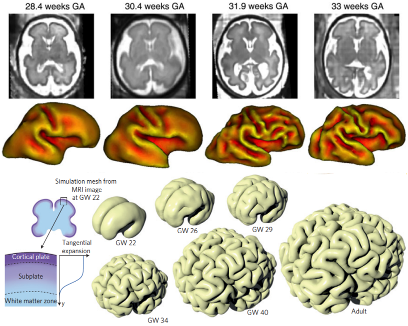
Bénéfices/risques de l'IA pour l'environnement : Que faire en contexte d'incertitude ?

Julien Lefèvre

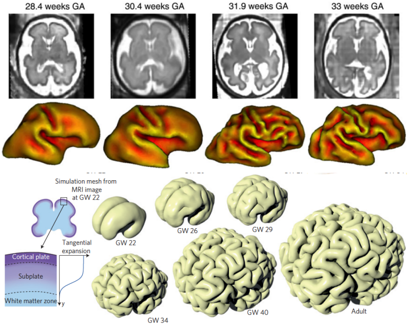
Aix-Marseille Université (AMU)
julien.lefevre@univ-amu.fr

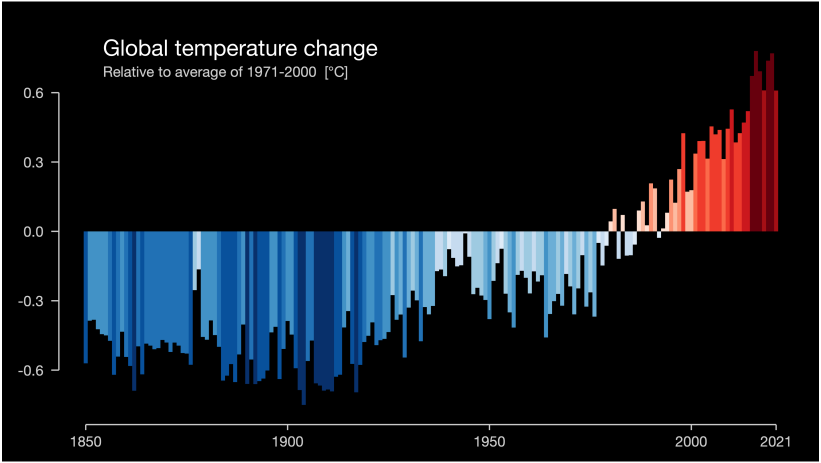
4 Avril 2023

Before and after the Satori



Before and after the Satori





https://showyourstripes.info/c/globe

sustainability MDPI

Article

Unraveling the Hidden Environmental Impacts of AI Solutions for Environment Life Cycle Assessment of AI Solutions

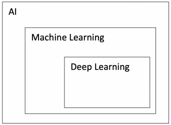
Anne-Laure Ligozat [1,*], Julien Lefevre [2], Aurélie Bugeau [3] and Jacques Combaz [4]

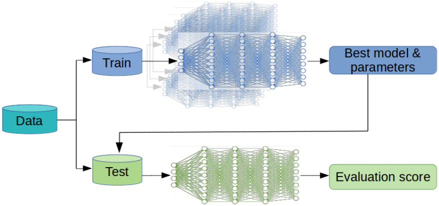
[1] Université Paris-Saclay, CNRS, ENSIIE, Laboratoire Interdisciplinaire des Sciences du Numérique, 91400 Orsay, France
[2] Aix Marseille Univ, CNRS, INT, Inst Neurosci Timone , Marseille, France; julien.lefevre@univ-amu.fr
[3] Univ. Bordeaux, CNRS, Bordeaux INP, LaBRI, UMR5800, F-33400 Talence, France aurelie.bugeau@u-bordeaux.fr
[4] Univ. Grenoble Alpes, CNRS, Grenoble INP, VERIMAG, 38000 Grenoble, France; jacques.combaz@univ-grenoble-alpes.fr
* Correspondence: anne-laure.ligozat@lisn.upsaclay.fr

1 - Context

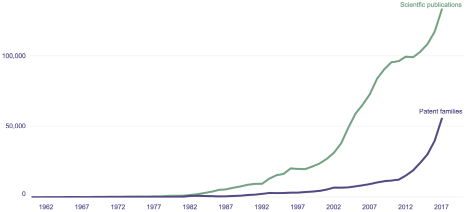
AI

Abuse of language





AI growth



[WIPO Technology Trends 2019]

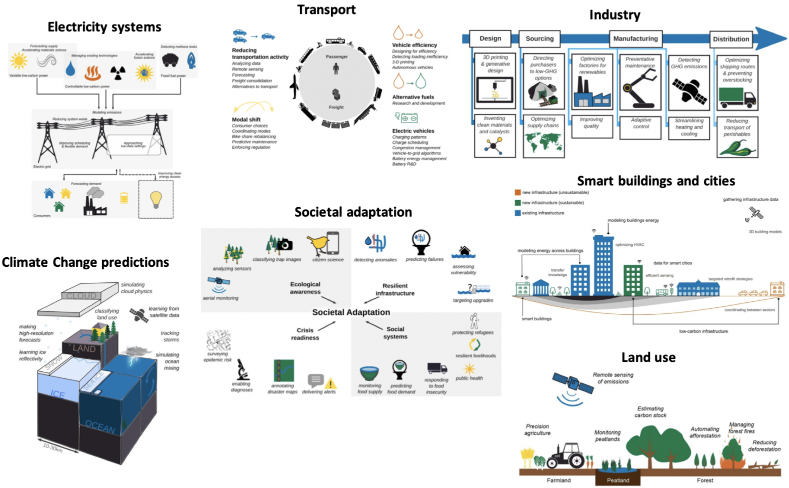
AI and environment

2 communities

- AI for green : AI as a tool for sustainable development
- green AI : measure and reduce AI impacts

Tackling climate change with machine learning [Rolnick et al., 2019]

826 references in 13 domains



Tackling climate change with machine learning [Rolnick et al., 2019]

Techno-optimist perspective.
But **uncertain impacts** and Jevon's paradox are mentioned :

- Transportation : *"autonomous vehicles could cause people to drive far more"*
- Industry : *"it is worth noting that greater efficiency may increase the production of goods" and thus GHG emissions"*
- Farms and forests : *"making forestry more efficient can have a negative effect by increasing the amount of wood harvested"*

Tackling climate change with machine learning [Rolnick et al., 2019]

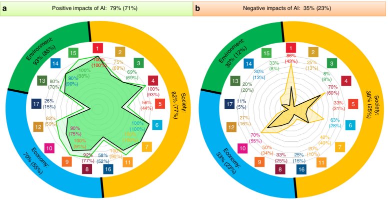
Recommendations are proposed :

"In all these cases, projects should be pursued with great care so as not to impede or prolong the transition to a low-carbon electricity system ; ideally, projects should be ***preceded by system impact analyses*** *to ensure that they will indeed decrease GHG emissions."* p10

"When designing and promoting new mobility services, it is important that industry and public policy prioritize lowering GHG emissions. Misaligned incentives in the early stages of technological development could result in the ***lock-in to a service with high GHG emissions****"* p15

Role of AI in achieving the Sustainable Development Goals

17 sustainable development goals



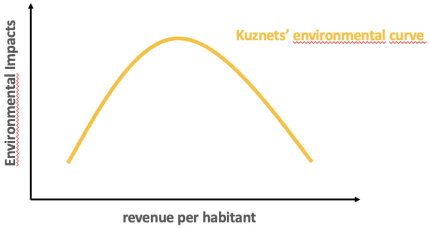
444 references for 169 goals

Some surprising aspects

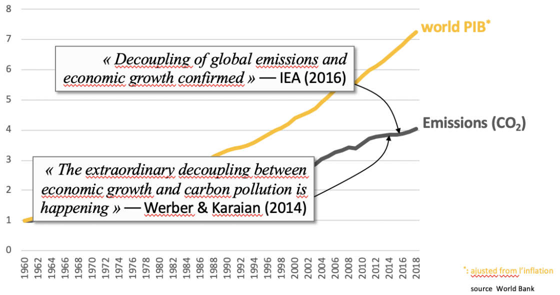
- agriculture : *"Blockchain technologies have been identified in the literature as offering potential to support vulnerable individuals and communities with affordable, secure and trusted documentation relating to rights or ownership over land"*
- energy efficiency : mostly everything is positive !

[Vinuesa et al., 2020]

Relation with narratives on decoupling



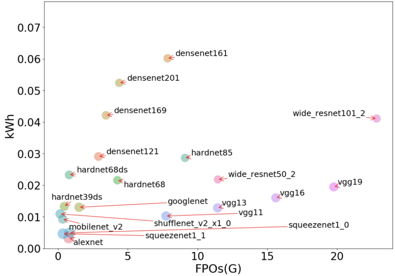
"Un peu de croissance pollue, beaucoup de croissance dépollue" - Laurence Parisot



Measuring impacts while training (green AI)

Method 1 : FPO (floating-point operations, i.e. number of basic operations)[Schwartz et al., 2020, Canziani et al., 2017, Henderson et al., 2020]

- Reflects the amount of calculations, Partially correlated with execution time, hardware independent
- Not completely correlated to energy consumption



Measuring impacts while training (green AI)

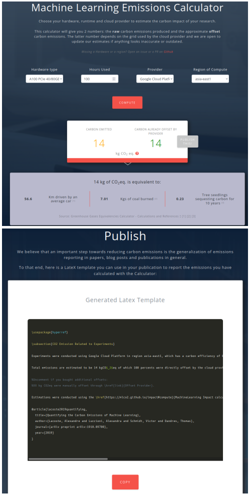
Method 2 : Measuring energy use
[Schwartz et al., 2020, Canziani et al., 2017]

$$E_{total} = \mathrm{PUE}\left(\mathrm{E_{dram}} + \mathrm{E_{cpu}} + \mathrm{E_{gpu}}\right)$$

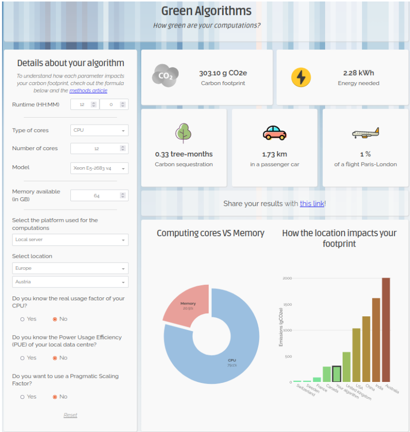
then convert it to CO2e

$$CO2e = EmissionFactor * E_{total}$$

Some existing tools



https://mlco2.github.io/impact/



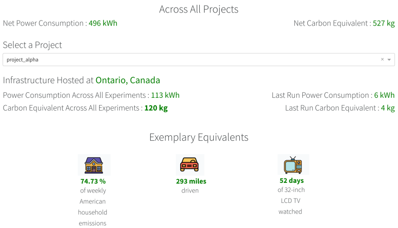
http://www.green-algorithms.org/

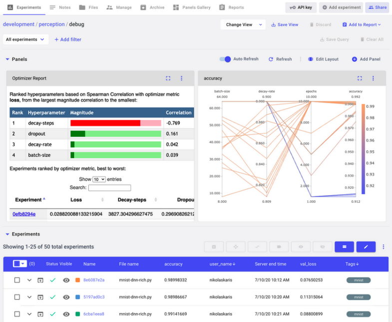
Some existing tools : code

codecarbon (mlco2)

```python
from codecarbon import EmissionsTracker

tracker = EmissionsTracker()
tracker.start()
# GPU Intensive code goes here
tracker.stop()
```





Some results

- 4 NLP models (natural language processing) [Strubell et al., 2019]
- Estimated energy consumption while training

Model	Hardware	Power (W)	Hours	kWh·PUE	CO_2e	Cloud compute cost
Transformer$_{base}$	P100x8	1415.78	12	27	26	\$41–\$140
Transformer$_{big}$	P100x8	1515.43	84	201	192	\$289–\$981
ELMo	P100x3	517.66	336	275	262	\$433–\$1472
BERT$_{base}$	V100x64	12,041.51	79	1507	1438	\$3751–\$12,571
BERT$_{base}$	TPUv2x16	—	96	—	—	\$2074–\$6912
NAS	P100x8	1515.43	274,120	656,347	626,155	\$942,973–\$3,201,722
NAS	TPUv2x1	—	32,623	—	—	\$44,055–\$146,848
GPT-2	TPUv3x32	—	168	—	—	\$12,902–\$43,008

lbs

Most widley used model at the time
652 kg C02e
~1 round-trip Paris Hong Kong by plane
or ~2 500km en voiture

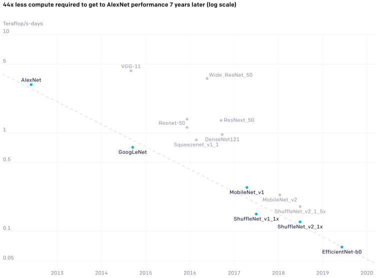
PUE = 1.58 (Ascierto, 2018)
FE = 0.954

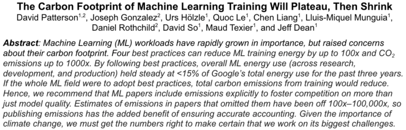
Conclusions from authors

- Need for cost-resource/accuracy analysis
- Public and shared computing resources for academics (price)
- Development of less consuming hardware and algorithms

Limits of existing works

- AI community not well aware of existing works on impacts
- Mono-criterion
- Perimeter : electricity consumption during training
- Promises





The Carbon Footprint of Machine Learning Training Will Plateau, Then Shrink

David Patterson[1,2], Joseph Gonzalez[2], Urs Hölzle[1], Quoc Le[1], Chen Liang[1], Lluis-Miquel Munguia[1], Daniel Rothchild[2], David So[1], Maud Texier[1], and Jeff Dean[1]

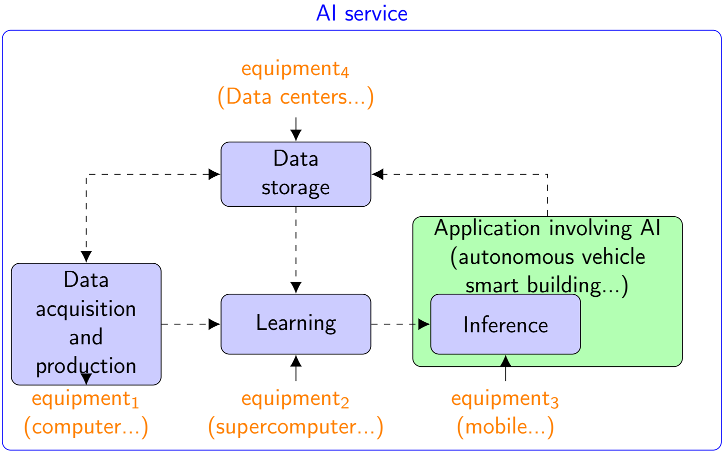
Abstract*: Machine Learning (ML) workloads have rapidly grown in importance, but raised concerns about their carbon footprint. Four best practices can reduce ML training energy by up to 100x and CO_2 emissions up to 1000x. By following best practices, overall ML energy use (across research, development, and production) held steady at <15% of Google's total energy use for the past three years. If the whole ML field were to adopt best practices, total carbon emissions from training would reduce. Hence, we recommend that ML papers include emissions explicitly to foster competition on more than just model quality. Estimates of emissions in papers that omitted them have been off 100x–100,000x, so publishing emissions has the added benefit of ensuring accurate accounting. Given the importance of climate change, we must get the numbers right to make certain that we work on its biggest challenges.*

Reducing AI impacts (green AI)

- Reduce number of parameters of neural networks, size of datasets, etc
- Frugal / sober AI → 24th-25th November in Grenoble
- Dismantle AI [Couillet 2022]

2 - How to evaluate AI services dedicated to environment ?

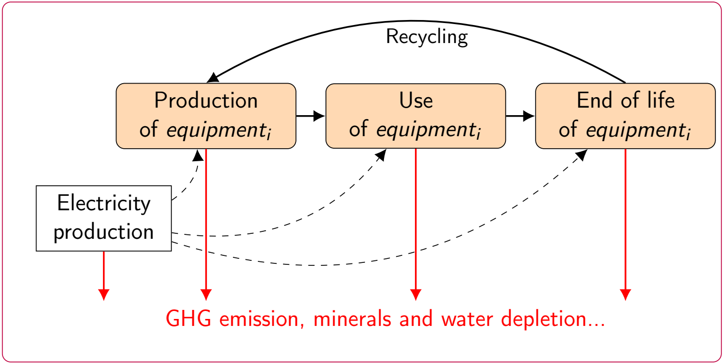
Lifecycle of an AI service



Environmental impacts of a service

The lifecycle assessment (LCA) of a service relies on :

- LCAs of the equipments/devices involved in the service
- allocation procedures (keys) to determine the fraction of the impacts attributed to the service.



Life cycle assessment (LCA) of an equipment

Environmental benefits of an AI solution

The (potential) benefits of an AI solution M_2 which is a substitute for a reference solution M_1 (typically without AI) are deduced from a comparative LCA :

$$\Delta(M_2|M_1) = LCA(M_2) - LCA(M_1) \in \mathbb{R}^d \tag{1}$$

Common pitfalls in such evaluations :

- focusing on a single environmental criterion (burden shift)
- restricted perimeter
- generalization of contextual solutions/benefits
- third-order effects (e.g. rebound effects) are overlooked.

Example : a smart-building

Commonly overlooked aspects :

- lifecycle impacts of devices (sensors, DC, networks, ...)
- impacts of the training phase
- what are the impacts of the smart solution(s) on the design of the building ?
- are claimed energy gains reliable in the operational context ?
- is the solution(s) applicable to old buildings ?
- human interactions : rebound effects, counteracting behaviors.

3 - Literature study

Is AI for green aware of its own impact ?

If yes, how is it measured ?
Do the authors are able to say that a new solution is better than an old one ?

		Causal inference	Computer vision	Interpretable models	NLP	RL & Control	Time-series analysis	Transfer learning	Uncertainty quantification	Unsupervised learning
Mitigation	Electricity systems									
	Enabling low-carbon electricity		•	•		•	•		•	•
	Reducing current-system impacts		•				•		•	•
	Ensuring global impact		•					•		•
	Transportation									
	Reducing transport activity		•				•		•	•
	Improving vehicle efficiency		•			•				
	Alternative fuels & electrification					•				•
	Modal shift	•	•				•		•	
	Buildings and cities									
	Optimizing buildings	•				•	•	•		
	Urban planning		•				•	•		•
	The future of cities				•			•	•	•
	Industry									
	Optimizing supply chains		•			•	•			
	Improving materials									•
	Production & energy		•	•		•				
	Farms & forests									
	Remote sensing of emissions		•							
	Precision agriculture		•			•	•			
	Monitoring peatlands		•							
	Managing forests		•			•	•			
	Carbon dioxide removal									
	Direct air capture									•
	Sequestering CO_2		•						•	•
Adaptation	Climate prediction									
	Uniting data, ML & climate science		•	•			•		•	
	Forecasting extreme events		•	•			•		•	
	Societal impacts									
	Ecology		•					•		
	Infrastructure					•	•		•	
	Social systems		•				•			•
	Crisis		•		•					
	Solar geoengineering									
	Understanding & improving aerosols						•		•	
	Engineering a control system					•			•	
	Modeling impacts						•		•	
Tools for Action	Individual action									
	Understanding personal footprint	•			•	•	•			
	Facilitating behavior change				•					•
	Collective decisions									
	Modeling social interactions			•		•				
	Informing policy	•	•		•				•	•
	Designing markets					•	•			•
	Education				•	•				
	Finance				•		•		•	

[Rolnick et al., 2019]

Case studies

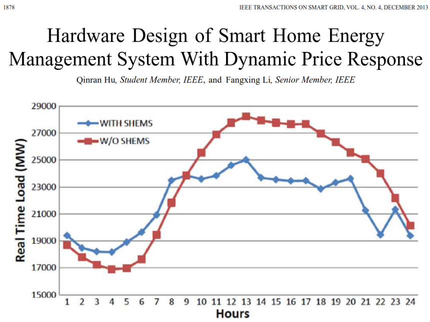
Analysis of several domains in [Rolnick et al., 2019] identified with "high leverage" (Modelling demand/freight, Electric vehicle, Low carbon, Smart buildings).

57 articles have an environmental evaluation following the categories :

a. No mention of the environmental gain.
b. General mention of the environmental gain.
c. No quantitative evaluation or only indirect estimation.
d. Evaluation of the energy gain without the AI program.
e. Evaluation of the energy gain taking the use phase of the AI service into account.
f. Comprehensive evaluation of the environmental gain (comparison of LCAs).

A few examples

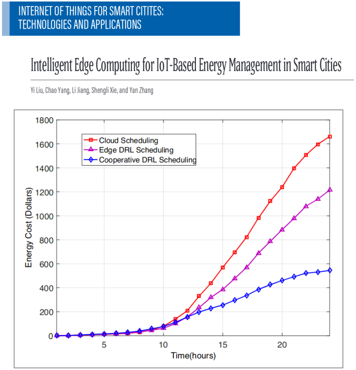
Guess which category ! :-)


1878 IEEE TRANSACTIONS ON SMART GRID, VOL. 4, NO. 4, DECEMBER 2013

Hardware Design of Smart Home Energy Management System With Dynamic Price Response

Qinran Hu, *Student Member, IEEE*, and Fangxing Li, *Senior Member, IEEE*

ML methods : Bayes classifier, Hiden Markov Model


INTERNET OF THINGS FOR SMART CITITES: TECHNOLOGIES AND APPLICATIONS

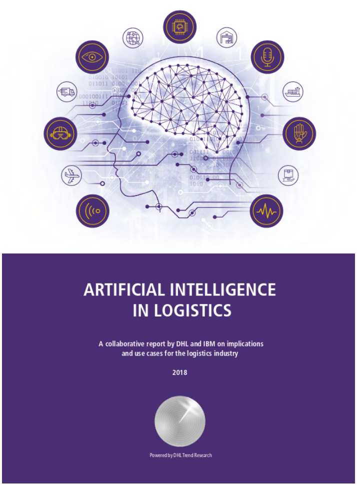
Intelligent Edge Computing for IoT-Based Energy Management in Smart Cities

Yi Liu, Chao Yang, Li Jiang, Shengli Xie, and Yan Zhang

Deep Learning

A few examples

Guess which category ! :-)



1.3 Trends Accelerating AI

Core technological advances are central to the continued development of AI. Significant progress has been made with all core AI technologies, and the levels of investment and demand for ongoing improvement give good reason to expect this growth will continue well into the future. Technological advances can be classified into three broad categories: improving computer processing speed and power, increasing AI system access to big data, and using algorithmic improvements to enable more complex AI applications.

Computing Power & Speed: AI is a computer processing-intensive technology – breakthroughs in computing power and efficiency have enabled the expansion and complexity of AI applications. In the technology industry, Moore's Law is used to show the relationship between the cost and speed of computer processing power over time, the trajectory of which results in an exponential curve as seen in figure 12. Until recently, a computing device's CPU, or central processing unit, typically provided the core function of processing. In recent years, GPUs, or graphical processing units, have begun to partially take over computer processing workloads, contributing significantly to the rise of AI.

Originally designed for the much larger and more complex computational workloads of rendering computer gaming graphics in real time, GPUs are designed to handle hundreds of tasks in parallel, and today are successfully being used to enable AI applications.

Advances in computer chip technology are an important part of the AI developmental story. Given the consistency of chip improvements and the likelihood that chip design will continue to improve, this is not the primary reason for the existence of AI but just one of the essential enablers.

THE DIFFERENCE BETWEEN A CPU AND GPU

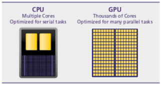


Figure 11: Comparison of CPUs and GPUs; **Source:** Nvidia

LAW OF ACCELERATING RETURNS

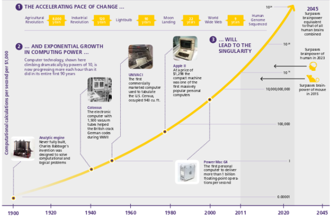


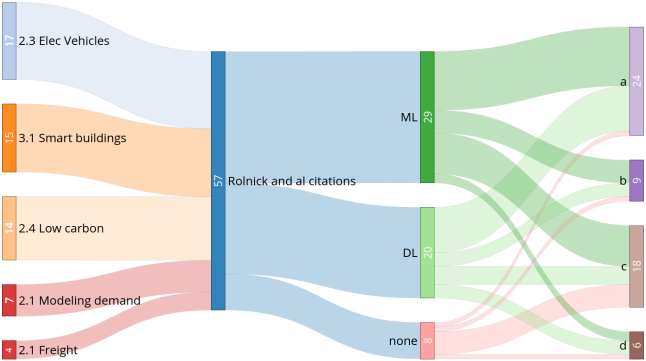
Figure 12: Ray Kurzweil's Law of Accelerating Returns depicts the exponential growth of computer processing power and technology innovations throughout history, and anticipates computers will exceed human intelligence in the future; **Source:** TIME / Wikipedia

Main result

We found no studies that follow the categories :

- e Evaluation of the energy gain taking the use phase of the AI service into account.
- f Comprehensive evaluation of the environmental gain (comparison of LCAs).

How to evaluate AI services dedicated to environment ?



- a No mention of the environmental gain.
- b General mention of the environmental gain.
- c No quantitative evaluation or only indirect estimation.
- d Evaluation of the energy gain without the AI program.

To summarize

- Framework to evaluate the direct impacts of AI for green applications
- The analysis of literature on AI has raised several issues

Discussion

- Current Environmental Evaluation of AI Services is Under-Estimated (mostly energy/GHG)
- Many unknowns in the LCA (e.g. production of GPU)
  => *AI community should lobby the companies to open a part of their data*
- Large structural changes (Third-order effects) are difficult to take into account e.g through the IoT whose large deployment may impact material demand (Lithium, Cobalt...)
  => *Consequential LCA instead of attributional LCA*
- Narratives about dematerialization and techno-solutionism
  => *Many scientists question techno-solutionism as the only answer to the environmental emergency.*
  => *The promises suggested by these new technologies should be more debated.*

4 - Remise en perspective

Que faire, en contexte d'incertitudes ?

"Que faire ?"

- Question récurrente, au niveau des individus, après une prise de conscience écologique.
- Question a priori moins habituelle dans l'ESR où créer et transmettre du savoir n'est pas toujours perçu comme relevant d'un "faire".
- Fondement même de la politique (*politikos/politeia*)

La question peut être abordée au travers des différent.e.s acteurs.rices impliqué.e.s : décideurs, industriels, scientifiques, citoyen.ne.s. (inspiration atelier SEnS)
On reste dans le cadre du numérique en élargissant un peu la perspective (risques sociétaux, démocratiques etc).

4 - Remise en perspective

Agir du point du/de la citoyen.ne

- Ecogestes
- Sobriété
- Boycott

Selon carbone4, les actions individuelles vertueuses pourraient conduire à une réduction d'un quart des émissions GES.

4 - Remise en perspective

Agir du point de vue des décideurs

- Rapports en amont, exemples : rapport Villani (2018), rapport IoT (2022)
- Autorités de régulation (ARCEP)
- Feuille de route du numérique (2021)
- Loi REEN (2021)

Innovation politique via la convention citoyenne pour le climat et le moratoire sur la 5G.

4 - Remise en perspective

Agir du point de vue des industriels

- Contraintes de rentabilité économique
- Ecueil du greenwashing, via le discours sur l'efficacité
- Lobbying hasardeux (exemple : le Metavers et ses bénéfices sur l'environnement)

Moratoire pour proposer une pause sur l'apprentissage d'IA géante.

4 - Remise en perspective

Agir du point de vue de la/du scientifique

- Utiliser ses compétences pour développer des projets de recherche en lien avec l'écologie (risque de greenwashing).
- Compte tenu de l'incertitude sur les impacts de l'IA/du numérique, aller vers plus d'études ou d'analyses pour permettre une meilleure quantification (labos1point5, ecoinfo).

Proposer des recommandations (exemple : document de cadrage Évaluation environnementale de projets impliquant des méthodes d'IA par ecoinfo)

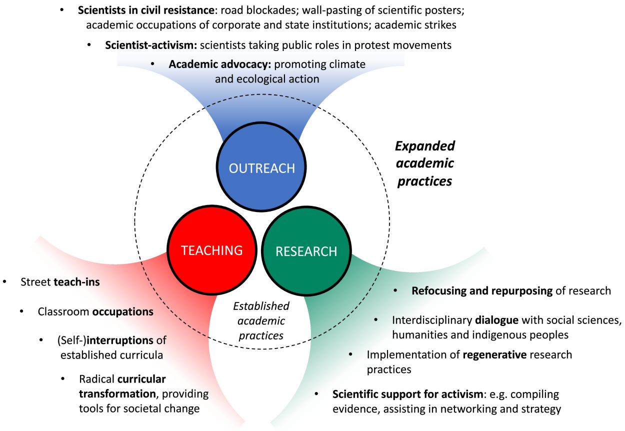
4 - Remise en perspective

Agir avec reflexivité

- La production, d'artefacts ou même de connaissances s'inscrit dans un tissu d'act.eur.rice.s, avec des dépendances, des rapports de force, des discours etc
- Prendre conscience de cet écosystème, l'étudier et en tirer un discours critique fait partie d'une activité réflexive
- C'est un des sujets d'intérêt des atecopols (Toulouse, Paris, Marseille etc)

En parallèle et en lien avec un renouveau de la pensée techno-critique des années 70 (Ellul, Illich, Gorz...) avec du sang neuf (F. Jarrige, J.B. Fressoz, A. Monnin, F. Lopez...)

Diversité des formes d'action



[Racimo et al., 2022]

Pour finir : Un projet réflexif ?

What Neuroscience in the Anthropocene era ?

Objectives

Global aim

Entrain the NeuroMarseille community to an environmentally responsible research

Specific aims

- redefining neuroscientific questions for <u>understanding the actual crisis</u> (WP1)
- <u>interdisciplinary meetings</u> on neuroscience at the era of anthropocene (WP2)
- adopting a <u>sociological vision</u> on shared values, unthought and wills for the <u>future of the community</u> with respect to the climate crisis (WP2)
- identifying, sharing and promoting existing <u>virtuous practices</u> and evaluation of the <u>impact</u> of neuroscientific research. Improving scientific practices (WP3).

- Endorsement by UNESCO: International Year of Basic Sciences for Sustainable Development.

References I

Canziani, A., Paszke, A., and Culurciello, E. (2017).
An Analysis of Deep Neural Network Models for Practical Applications.
arXiv :1605.07678.

Henderson, P., Hu, J., Romoff, J., Brunskill, E., Jurafsky, D., and Pineau, J. (2020).
Towards the systematic reporting of the energy and carbon footprints of machine learning.
Journal of Machine Learning Research, 21(248) :1–43.

Racimo, F., Valentini, E., Rijo De León, G., Santos, T. L., Norberg, A., Atmore, L. M., Murray, M., Hakala, S. M., Olsen, F. A., Gardner, C. J., et al. (2022).
The biospheric emergency calls for scientists to change tactics.
Elife, 11 :e83292.

Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K., Ross, A. S., Milojevic-Dupont, N., Jaques, N., Waldman-Brown, A., Luccioni, A., Maharaj, T., Sherwin, E. D., Mukkavilli, S. K., Kording, K. P., Gomes, C., Ng, A. Y., Hassabis, D., Platt, J. C., Creutzig, F., Chayes, J., and Bengio, Y. (2019).
Tackling Climate Change with Machine Learning.
arXiv :1906.05433 [cs, stat].
arXiv : 1906.05433.

Schwartz, R., Dodge, J., Smith, N. A., and Etzioni, O. (2020).
Green AI.
Communications of the ACM, 63(12) :54–63.

References II

Strubell, E., Ganesh, A., and McCallum, A. (2019).
Energy and policy considerations for deep learning in NLP.
In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 3645–3650, Florence, Italy. Association for Computational Linguistics.

Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., Felländer, A., Langhans, S. D., Tegmark, M., and Fuso Nerini, F. (2020).
The role of artificial intelligence in achieving the Sustainable Development Goals.
Nature Communications, 11(1) :233.

Recent initiatives of scientists

Appel de 1000 scientifiques

« Face à la crise écologique, la rébellion est nécessaire »

Table ronde XR pendant l'UICN

2018 — 2019 — 2020 — 2021 — 2022

- *World Scientists' Warning to Humanity: A Second Notice*, Ripple et al + 15.000 scientists
- Création Atecopol Toulouse
- Labos1point5
- Pétition => Conseil du Climat d'AMU
- Création Atecopolam

Face à l'urgence climatique, les scientifiques doivent réduire leur impact sur l'environnement